Calhoun: The NPS Institutional Archive
DSpace Repository

Theses and Dissertations 1. Thesis and Dissertation Collection, all items

1981

Evaluation of the artificial intelligence program STAMMER2 in the tactical situation assessment problem.

Ferranti, John Peter

Monterey, California. Naval Postgraduate School

http://hdl.handle.net/10945/20521

Downloaded from NPS Archive: Calhoun

Calhoun is the Naval Postgraduate School's public access digital repository for research materials and institutional publications created by the NPS community. Calhoun is named for Professor of Mathematics Guy K. Calhoun, NPS's first appointed -- and published -- scholarly author.

Dudley Knox Library / Naval Postgraduate School
411 Dyer Road / 1 University Circle
Monterey, California USA 93943

http://www.nps.edu/library

NAVAL POSTGRADUATE SCHOOL

Monterey, California

THESIS

EVALUATION OF THE ARTIFICIAL INTELLIGENCE PROGRAM STAMMER2 IN THE TACTICAL SITUATION ASSESSMENT PROBLEM

by

John Peter Ferranti Jr.

March 1981

Thesis Advisor: D. R. Barr
G. T. Howard

Thesis
F2643

Approved for public release, distribution unlimited.

NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

EVALUATION OF THE ARTIFICIAL INTELLIGENCE PROGRAM STAMMER2 IN THE TACTICAL SITUATION ASSESSMENT PROBLEM

by

John Peter Ferranti Jr.

March 1981

Thesis Advisor: D. R. Barr
G. T. Howard

Approved for public release, distribution unlimited.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Evaluation of the Artificial Intelligence Program Stammer 2 in the Tactical Situation Assessment Problem		5. TYPE OF REPORT & PERIOD COVERED Master's Thesis; March 1981
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) John Peter Ferranti, Jr.		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		12. REPORT DATE March 1981
		13. NUMBER OF PAGES 89
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

STAMMER2
Tactical Situation Assessment (TSA)

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

STAMMER2 (System for Tactical Assessment of Multisource Messages, Even Radar) is an experimental program created as part of an investigation into methods of correlating information in the naval environment. This thesis is an exploration into the application of artificial intelligence to the tactical situation assessment problem and into various evaluation methodologies for STAMMER2. Included is an overview of one of these

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE
S/N 0102-014-6601 |

experiments, using the facilities of the Naval Postgraduate School Command, Control and Communications Laboratory and the Naval Ocean Systems Center, San Diego.

Approved for public release, distribution unlimited.

Evaluation of the Artificial Intelligence Program STAMMER2 in the Tactical Situation Assessment Problem

by

John Peter Ferranti,Jr.
Lieutenant Commander, United States Navy
B.S., Marquette University, 1968

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN SYSTEMS TECHNOLOGY - C3

from the

NAVAL POSTGRADUATE SCHOOL
March 1981

ABSTRACT

STAMMER2 (System for Tactical Assessment of Multisource Messages, Even Radar) is an experimental program created as part of an investigation into methods of correlating information in the naval environment. This thesis is an exploration into the application of artificial intelligence to the tactical situation assessment problem and into various evaluation methodologies for STAMMER2. Included is an overview of one of these experiments, using the facilities of the Naval Postgraduate School Command, Control and Communications Laboratory and the Naval Ocean Systems Center, San Diego, California.

TABLE OF CONTENTS

I. INTRODUCTION ---------- 9

A. BACKGROUND ---------- 9

B. OBJECTIVES ---------- 10

II. ARTIFICIAL INTELLIGENCE ---------- 11

A. DEFINITION ---------- 11

B. DISCUSSION ---------- 11

III. TACTICAL SITUATION ASSESSMENT AND STAMMER2 ---------- 16

A. TACTICAL SITUATION ASSESSMENT ---------- 16

B. STAMMER2 ---------- 18

IV. EXPERIMENTAL METHODOLOGY ---------- 30

A. STAGE-ONE EXPERIMENTS ---------- 34

1. Objective ---------- 34

2. Resources Required ---------- 34

3. General Context ---------- 35

4. Evaluation ---------- 36

5. Comments and Special Instructions ---------- 37

B. STAGE-TWO EXPERIMENTS ---------- 38

1. Objective ---------- 38

2. Resources Required ---------- 38

3. General Context ---------- 40

4. Evaluation ---------- 41

5. Comments and Special Instructions ---------- 42

C. STAGE-THREE EXPERIMENTS ____________________ 43

1. Objective ______________________________ 43

2. Resources Required _____________________ 43

3. General Context ________________________ 45

4. Evaluation _____________________________ 46

5. Comments and Special Instructions ________ 48

V. SAMPLE EXPERIMENT ______________________________ 50

A. DESCRIPTION ___________________________________ 50

B. RESULTS AND CONCLUSIONS ______________________ 54

APPENDIX A STAMMER2 GUIDE __________________________ 67

APPENDIX B DATA BASE _______________________________ 69

APPENDIX C RULES ___________________________________ 71

APPENDIX D MESSAGES ________________________________ 82

APPENDIX E SCENARIO ________________________________ 84

LIST OF REFERENCES _________________________________ 86

INITIAL DISTRIBUTION LIST __________________________ 88

LIST OF FIGURES

1. STAMMER2 Graphics Display, Game Minute 3 __________59

2. STAMMER2 Graphics Display, Game Minute 14 _________60

3. STAMMER2 Graphics Display, Game Minute 45 _________61

ACKNOWLEDGEMENTS

I wish to acknowledge the patient instruction of my thesis advisors, Professor Donald R. Barr and Professor Gilbert T. Howard, who have made this thesis research an enjoyable and educational experience. Additionally, I would like to note the time and assistance given by Mssrs. Dennis C. McCall and Robert J. Bechtel of Naval Ocean Systems Center, San Diego, California, whose technical expertise made the task possible. The primary acknowledgement is to my wife, Anne, who gave me the support and inspiration for my work.

I. INTRODUCTION

A. BACKGROUND

STAMMER2 (System for Tactical Assessment of Multisource Messages, Even Radar) is an experimental computer program created as part of an investigation into new methods of correlating information available in a naval environment in response to the tactical situation assessment (TSA) problem. It was written by the Tactical Command and Control Division of the Naval Ocean Systems Center, San Diego, California. STAMMER2 has served as a testbed for explorations of applications of rule-based artificial intelligence inference systems, with a general and flexible approach to the range of acceptable inputs to the system.

In practical application, information in a basic format is received from a variety of sensors. This information is processed within a network of interconnected rules or conditional statements which may be satisfied by this information, even responding to information in cases involving varying degrees of uncertainty. STAMMER2 is capable of performing deductions based on this received information to speed the decision process in a rapidly changing environment.

Before we discuss the specifics of STAMMER2 and a method of testing the utility of this decision aid, background information will be presented to set the stage for the why, where, and how such a system was developed.

B. OBJECTIVES

There are three principal objectives for this thesis:

1. Identify the characteristics of an artificial intelligence program, with an overview of the applications to a specific military situation.
2. Propose various levels of evaluation methodologies for STAMMER2.
3. Present the results of an experiment conducted with STAMMER2 using one of the methodologies presented.

II. ARTIFICIAL INTELLIGENCE

A. DEFINITION

Artificial intelligence (AI) is defined as "the study of ideas which enable computers to do the things that make people seem intelligent" [Ref. 1] and "the science of making machines do things that would require intelligence if done by men" [Ref. 2].

The intricate processes from which these "things" have evolved, such as learning, growth, or maturation are bypassed in the artificial intelligence program to create a facsimile of human thought and, subsequently, decision-making. But the bypassing is, in reality, a very careful compilation of the elements which are, or seem to be, parts of the human thought patterns and strategies leading to decisions. That is, human thought is analyzed and dissected into discrete elements which may be recombined artificially to mimic human intelligence. Alternatively, differing patterns or methods are explored to uncover new routes which could, in fact, lead to similar results.

B. DISCUSSION

To imagine the thought process as a branched tree, with decisions made at each juncture, or node,

is a simplification which does not reflect the complexity of the problem of describing human thought. The actual representation must consider many interconnected paths, with multidirectional flow through those paths. A control mechanism must also be considered which carries a thought through to conclusion and minimizes retraced or non-productive routes.

A goal of artificial intelligence inquiry is to duplicate this complex structure in a computer program. A difficulty lies in the multi-dimensional nature of the thought process which requires equally complex investigation. By examining the problem from the viewpoint of discrete elements, it may be possible to find the combination which accurately reflects a pattern which not only duplicates what happens in the case of a specific problem, but also a general case which may be applicable to other problems.

If STAMMER2 can duplicate the reasoning process of a decision maker, with a data base and memory consistent with the situation presented, then it is possible to use this artificial intelligence program as an "intelligent assistant" in the tactical situation assessment problem which will be discussed in the next section.

The computer is the "ideal experimental animal." Infinitely patient and easy to care for, the computer requires structured instruction, composed of discrete

elements, which it combines as programmed. The combination and recombination, done at speeds sufficient to expose unproductive paths rapidly, can provide a practical test of the precision with which the experimenter defines his thought model [Ref. 3]. Additions or deletions which the experimenter may hypothesize as crucial (or useless) to the final result of his pattern may be evaluated with relative ease. Flaws or conceptual mistakes which may exist are likely to be detected in the computer because the computer generally will not accept ambiguity of instruction and may simply cease execution of the program. However, a program designed to answer a certain sort of question does not guarantee a correct answer [Ref. 4].

The applicability of computers and programming is not the emphasis of artificial intelligence. However, the unambiguous nature of the program is mentioned to reinforce the " transparency" of the actual machine and focus on the precision of understanding human thought.

Using artificial intelligence techniques, research has been done to explore vision, problem solving, language, and neurosis [Ref. 5]. Some of these models have been successful in predicting behavior under specified conditions while others have been at least of metaphorical value in helping to understand the possible mechanism of intelligence. By representing thought and knowledge, in

general, in a useful and flexible manner it is also possible to arrange an interaction between a human and the machine in a task oriented environment.

Knowledge is defined as a collection of facts, the state of knowing, or all that has been perceived by the human mind [Ref. 6]. Knowledge may be in human or machine memory or in a data base. It is, however, also in the actual procedures that operate on or by reference to the data. To store disjointed facts in computer memory or in a data base from which the labelled information units may be retrieved is not the goal of artificial intelligence. The identification of the desired fact (or a location code by which that fact may be retrieved) is a problem of database management. To discern which fact is required is, however, the end of a series of steps. In the problem solving case, the solution process is the strategy of gaining information after " a forward or backward looking reasoning method" [Ref 7].

Analogies to other human processes exist. For example, radar provides target information in a manner analogous to the eye. Data, such as target presence, range, and bearing may be useful to the operator as well as in a form which can be compared to characteristics stored in computer memory.

The specific goal of STAMMER2 is to practically combine this sensory data with the problem solving process of

data receipt, comparison and retrieval using the knowledge stored in the artificial intelligence program.

III. TACTICAL SITUATION ASSESSMENT AND STAMMER2

A. TACTICAL SITUATION ASSESSMENT

In the naval context, the ability of a commander to accurately assess his military situation requires decisions about his environment. While all decisions need not be made on an immediate basis, certain of them, called tactical, determine his immediate course of action. In an extreme example, the decision to launch a defensive weapon is based on a judgment that such an action is required based on the commander's evaluation. A statement of the tactical situation assessment (TSA) problem includes consideration of the information available to the decision maker.

In order to assess a military problem, the evaluator draws upon experience and training, after having considered the situation presented to him. The assessment of the tactical situation is, however, highly dependent upon factors associated with the decision maker himself.

The experience upon which a military decision maker draws is unique to him alone, although there is certainly a similarity of career patterns and situations common to many individuals. Unique to that individual is the quality of his recollection, his emotional or mental state, the perspective of the individual, and the length

of time elapsed since those events from which the experience was gained.

The actual decision maker in the operations center (the Combat Information Center or CIC) of a Navy ship is the Tactical Action Officer (TAO). During the period of the Vietnam conflict, the role of the Captain of a ship as final fighting authority was altered somewhat from the traditional role. The Tactical Action Officer concept was proposed, implemented in Navy Regulations, and exercised in combat [Ref. 8]. The concept basicly states that a trusted subordinate, trained and drilled in defensive doctrine and procedures, acts as weapons release authority in the temporary absence of the Captain.

The specialized formal training for TAOs, having as prerequisites proven maturity and operational experience, lasts approximately six weeks. The student is exposed to information concerning U.S. weapons systems and those of potential enemies. After this is committed to memory, "standard" tactical doctrine is exercised in a series of increasingly complex scenarios. The scenarios are presented in sequences allowing questions to be answered based on information normally available to the TAO (from the training, experience, tactical publications,sensors, and intelligence sources).

The TAO becomes the focus of the Combat Information Center. All sensor reports are displayed to him, all

messages from external sources are given to him directly or in summarized form. He can stand alone, on his own ship, or be electronically linked by radio to TAOs on other ships, each going through similar procedures and actions. This defines a sphere of knowledge in which the TAO operates, and can generally be delineated by the range of his sensors. A correlation function is required when the TAO receives "external" reports from sensors or platforms other than his own--in effect from a sphere of larger radius.

Assumptions involved in reducing many TAO activities to computer form include: a) the basic, repetitive mass of information can be reduced to a computerized database; b) standardized procedures for data retrieval and correlation can be similarly reduced; c)the sensor/intelligence information can be presented in a standard format; and d) the thought process by which the TAO functions may be artificially reproduced.

B. STAMMER2

STAMMER2 is a revised version of STAMMER, a System for Tactical Assessment of Multisource Messages, Even Radar. It was created as part of an investigation of correlation methodologies, and served as a testbed for explorations of applications of rule-based inference systems to the tactical situation assessment (TSA) problem. STAMMER concentrated on

the specific task of merchant ship discrimination from all contacts reported by radar and external messages [Ref. 9].

STAMMER2 is an organizer of information. It collects information by receiving messages and sensor reports (radar, electronic support measures, and sonar), and organizes this raw data into graphic displays and textual commentary to aid in tactical situation assessment. Through the use of specified rules, the system combines this information to draw conclusions about the situation in the vicinity of the home ship. These combinations are reflected in both the STAMMER2 display and commentary. The system is available for the examination of raw data as well as information about why and how the conclusions were reached.

STAMMER2 deals with information on a real-time basis. As a message is received from a sensor or information source in STAMMER2-readable format, it becomes part of the data base on which the rules operate. This process is complicated by the fact that the information is suspect, and its arrival may not be in chronological order.

Although STAMMER2 was developed in response to Navy- sponsored investigation of the TSA problem, the U.S. Army has conducted similar investigations in the field of military intelligence analysis [Ref. 10]. Based upon a detailed study of the analysts' role, methods, and thought processes in intelligence production,

it may be possible to streamline intelligence analysis. In essence, this process used by the analyst is the methodology from which the AI program might be developed after a descriptive model is organized and exercised.

The manner in which STAMMER2 functions may best be described by analogy to the medical diagnosis system called MYCIN [Ref. 11]. MYCIN is a problem solving system by which a physician, responding to queries from the program, receives assistance in identifying and treating blood bacterial infections. The basis for the expertise resident in MYCIN is the knowledge obtained from interviewing physicians and experts and stored in over 300 "productions". Productions are memory structures on which the program operates and are in the following format.

If the infection type is primary-bacteremia,
the suspected entry point is the gastrointestinal tract,
and the site of the culture is one of the sterile sites,

then there is evidence that the organism is bacteroides.

Conceptually, inference rules are very simple. Every rule attempts to retrieve information from memory and, if it succeeds, constructs a new assertion in memory. If all the conditions are matched, then the rules are said to "fire". That is, all the actions which are defined in the rule are

then carried out. These actions may be modifications to the data base, assigning of confidences, and printing of formatted descriptions.

In a simple rule interpreter, the system maintains a list of rules, each with conditions which pertain to a central data base, which is constantly being updated. At regular intervals the system attempts to fire each rule; that is, it checks to see if the rule conditions are currently satisfied in the data base. If one of many conditions is not satisfied, a rule may fail to fire and, in this case, any partial work done in satisfying the rule is lost. In addition, each rule is retried at the time interval without regard to the nature of the intervening changes to the data base, duplicating results where no changes to the data base have occurred.

Because medical diagnosis, even by experts, involves varying degrees of uncertainty, MYCIN is written with a so-called certainty factor for each conclusion. The physicians upon whose information the productions were established were also asked to provide "strength values", or probabilities of accuracy and confidence, for their assertions. This introduces the ability to deal with uncertain evidence into the artificial intelligence system.

The productions used to reach a conclusion are part of the conclusion set. MYCIN answers questions about how and

why a fact was established or used. The production is recalled (by name or number) to show an enumeration of the facts presented in the premise set. General facts can be questioned without resorting to the entire diagnosis because each production "stands alone", to be actively examined or questioned. In this case the steps taken to arrive at that particular fact can also be traced. This tracing may act as a training aid for the user in addition to presenting alternatives which can be explored.

The critical aspects of STAMMER2's design may be divided into four parts. These are memory (the data base), rule interpretation, explanation, and graphics.

The rules of STAMMER2 consist of conditions, actions and confidences, with conditions and actions in the same structure as assertions with allowances made for variables which are "bound" to a rule by a binding function in LISP, the language in which STAMMER2 is written [Ref. 12]. Those variables which are applied to the rule conditions will be evaluated to see whether the condition succeeds or fails. If it fails, a "suspension" is created which corresponds to the remainder of the rule. As more information is added to the data base, those suspensions which can use the new data are revived and continue as before (either completing or suspending again). A suspension contains not only the remaining part of the rule, but also the bindings established by already satisfied conditions, if any. Even

when a condition succeeds, there may be other ways for it to be satisfied, so a suspension is left behind.

The individual rules are expressed in the following form.

```
RULE-NAME
(CONDITIONS
 (( <condition 1> )
  ( <condition 2> )
    .
    .
  ( <condition n> ))
 ACTIONS
 (( <action 1> )
  ( <action 2> )
    .
    .
  ( <action m> )
 CONF <a number between -1 and +1>
 PRINFORM
"This is a sentence describing the rule.")
```

STAMMER2 deals with information on a real-time basis. However, this process is complicated by the fact that the arrival of reports may not be in chronological order, with later information superseding or negating earlier reports. A "data stream" is used to bind information as it is received to the rule condition which it satisfies.

A data stream may be defined as a sequence of values, existing over time in a computation. If a program is executed in a conventional language, then the history of successive values of a variable forms a stream. Thus, in contrast to static data structures such as lists and arrays,

where all the elements exist at one time, streams are dynamic data structures, and are addressable objects in LISP. Only the new data received need be compared to the rules. New matches are added to the end of the stream. A temporary "freezing" of the action specified by an assertion occurs until after the new data is read onto the appropriate streams. The stream, then, allows the rapid review of stored information without the penalty of complete review of potentially irrelevant data.

In STAMMER2, confidence is provided in the rules by the creator of the rule. It is calculated dynamically on request at the time when it is displayed to the operator following a rule match and is not stored. Each assertion will have as its confidence the combination of the confidences of the rules and assertions which offer evidence for it. A confidence calculation is highly dependent upon the connective through which the rules and assertions are combined. An AND connective will, for example, display the confidence of the smallest value of all the rules combined. That is, the conclusion may never be stronger than the weakest piece of supporting evidence.

During execution of STAMMER2, the user will see the following cycle repeated as long as messages and reports are received into the message input file:

1. A message report is received, the user is informed, and the critical information in the message is printed for reference.

2. A display , showing the area situation with the new information, is drawn. The user may manipulate this image.

3. The system makes some commentary on the conclusions it can reach, on the basis of the new information.

4. If any conclusions were reached, the user is given the opportunity to query the system about the contents of its data base.

The message text is in the basic format of the received message. A contact name (if available), is displayed, followed by location and time of message. The user may pause to examine this message or select the graphic display mode.

Graphical support for STAMMER2 is provided by DSPLA, a software package developed at the Naval Ocean System Center specifically for tactical situation assessment [Ref. 13]. The DSPLA system is a collection of FORTRAN subroutines that allow storage, retrieval, and display of ship and aircraft tracks. This display, with the capability to show maps with latitude and longitude, may be manipulated by the user to vary the scale as desired. As soon as a message is received which contains a demonstrable assertion (such as a contact at a specific location) the display will appear on the display screen (if a display is available and selected in the initialization process). Examples of basic function keys are are listed below.

M	magnify about the cursor by a factor of 2
R	reduce about the cursor by a factor of 2

V center the view about the cursor position

; set type size to smallest on the TEKTRONIX 4214

C return to command mode

More detailed functions exist, with the full listing in NOSC Technical Document 252. The user may return to the command level to begin the explanation and query procedure.

The explanation system provides two primary capabilities: retrieval of memory and inference tracing. Although there would appear to be little difficulty in retrieving memory contents and tracing derivations, the display of this information in a human readable format is a major consideration. To make the user interface as natural as possible, the explanation system provides a query language that is "English-like". This language is an extremely limited version of English, which includes only certain types of questions and methods of phrasing those questions. However, the language was designed so that, while limited, it is sufficient to cover the user's needs without making its shortcomings apparent. The LISP function ASKUSER, which features recognition and prompting, is combined with a "prettyprinter" (to add words like "is" and "of" to memory contents) to simplify the user interaction.

A summary of the queries which the STAMMER2 user can ask, with general examples, follows. Following the initial query word, only those inputs shown will be allowed by the ASKUSER function. The connective words (in lower

case) will be provided by the "prettyprinter" [Ref. 14]. The words in parentheses may be inserted by the user as desired.

WHAT	Format: WHAT is (THE AN A) <RELATION> (OF) <ITEM> Example:WHAT is the COURSE of SIGHTING3 This query provides one of several question formats for asking about entries in the data base.
IS	Format: IS (THE) <ITEM> (A AN THE)<RELATION> (OF) <ITEM> Example:IS RADAR the SOURCE of SIGHTING3 This question form allows the user to ask about specific entries in the data base.
TELL me about	Format: TELL me about <ANYTHING> Example: TELL me about SIGHTING5 This is the most flexible query, and allows the user to ask rules, categories, relations, or specific facts.
WHERE	Format: WHERE is <OBJECT> or WHERE was <OBJECT> at <TIME> Example: WHERE is CONTACT7 With this command the user can ask about the position of a platform, merchant lane, or

storm at the present time, or in the future or past.

WHY

Format: WHY is <ASSERTION>

Example: WHY is A00345

With this command the user can find out the primary or immediate reasons that STAMMER2 used to conclude any assertion. All the rules involved in the decision will be displayed.

HOW

Format: HOW does rule <RULE> apply to <ASSERTION>

Example: HOW does rule ID-LANE apply to A234

This query allows the user to find out what assertions or facts were involved in permitting the rule to help conclude the given assertion.

WHOSE

Format: WHOSE <RELATION> is <ITEM>

Example: WHOSE TYPE is MERCHANT

This query acts as a partial inverse to the WHAT query.

WHO

Format: WHO is (THE AN A) <RELATION> (OF) <ITEM>

Example: WHO is INSIDE LANE1

This provides another form for querying the data base.

IV. EXPERIMENTAL METHODOLOGY

This section presents a discussion of experimental methodology which may be applied in the evaluation of STAMMER2. Following an overview of general issues, and a discussion of general levels of experimentation, three specific stages of experiment are presented. From these, one has been chosen to be carried out as a sample experiment.

The design of an experiment is based on the question to be addressed. An experiment to compare systems with different speeds of message receipt, for example, will include a message generator, an interval for testing, and the means to measure differences in total messages received. The criteria for "better" may be the higher number of messages received in a certain time or the shorter time needed to receive a certain number of messages. These objective measures of performance are established prior to the experiment, with sufficient flexibility to present the data to satisfy the question.

In more complex evaluation, the questions to be answered may be difficult to define. If the goal of the experiment is to find the number of messages received and understood during a definite time interval, the criteria for understanding must also become a part of the experimental procedure. Interpretation of the results of such an

experiment then becomes dependent on an "acceptable" level of understanding, an acceptable number of messages, or a combination of the two. This may then be an area where the objective measures give way to less precise but equally useful subjective evaluation.

In order to test the suitability of an artificial intelligence system, measures of performance may be difficult to establish. The components of the program may be individually examined (the structure, execution, results), or the overall "improvement" in a parameter of the function which the program is designed to assist may be tested. However, to show that the system would respond to a question or situation as quickly as a human decision maker, given that the situation were constrained to what the program and the human know and the information flow was at "normal" speed, may prove nothing more than that computers are faster than humans. Considering the difficulty in establishing meaningful measures of effectiveness, subjective evaluation of the capability and the utility of STAMMER2, by tactically oriented individuals, may show the greatest promise as an evaluation technique.

Experimentation may be performed in stages ranging from simple static debug/validation experiments, which concentrate on the capabilities of the technology and require few resources, to more extensive technical

evaluations of STAMMER2 with a dynamically changing data base in a controlled environment. This could include dynamic operational evaluations in simulated command and control environments pitting opposing forces against each other under various scenarios using one of the warfare environment simulator (WES) war game programs [Ref. 15].

Following an overview of three possible levels of testing, more detailed procedures for these levels will be presented.

The basic level of experiment consists of inputing into a data base various static threat scenarios (a snapshot view of the situation) concerning the status of "Blue" and "Orange" task forces. The test objective is to evaluate the capabilities of STAMMER2 to read the data base, signal the existence of the threats, and form opinions (based on specific rules selected by the test conductor) concerning other information in the form of sensor and intelligence reports.

The next level of experiment is conducted in a more realistic command and control environment with a dynamically changing data base. A detailed scenario is presented via WES (or its variants), presenting a wide range of inputs to STAMMER2 under changing circumstances to probe the limits of its capabilities, e.g., what threat intensity is required to "overload" STAMMER2, or what is

the length of time from the instant the report is received until STAMMER2 signals a conclusion.

In the most complex level of testing, WES is again used with Orange and Blue task forces. These forces are composed as closely as possible, within the capabilities of available test resources, with equipment (sensors, weapons, etc) and platforms that represent current military inventories. Experimental emphasis is shifted from a technical evaluation to an operational evaluation of the value of STAMMER2 to the decision maker. The objective is to determine if STAMMER2 improves the commander's ability to make rapid and accurate decisions. Operationally realistic scenarios are the desired norm. Each scenario should be replicated with different players with some trials having the Blue forces operating with STAMMER2 and some without. Operational measures of effectiveness, which depend on the scenarios under play, are selected to assess the operational utility of STAMMER2. In addition, the players and the umpire team are subjected to post-exercise interviews and questionnaires to get subjective evaluations of the worth of STAMMER2.

While there are no definitive boundaries between the levels of the formal experiments, the range of conditions must be designed for scenarios that are both realistic and informative. The actual analysis at each level of experiment must take into account the limitations of the scenario

generator [Ref. 16]. The three proposed levels of experiment will now be presented in detail.

A. STAGE-ONE EXPERIMENTS

1. Objective

Stage-one experiments evaluate the capability of STAMMER2 to identify, in a static environment, the tactical situation that confronts an afloat task force and to describe the situation semantically and graphically. This type of experiment establishes the type of threats that STAMMER2 can evaluate and assesses the timeliness of the warnings. The experiments may also serve to provide feedback through which STAMMER2 may be refined.

2. Resources Required

a. Technical

The STAMMER2 program and associated computer hardware are required for this experiment. At present, STAMMER2 is resident on the TOPS20 system at NOSC, San Diego, California.

b. Display

An alphanumeric display system will be required for players and the umpire team. In addition, a graphical display of the situation, as is normally generated by WES must be available to compare it with that presented by STAMMER2.

c. Scenario

The tactical scenario is presented as a series of operationally realistic "data plates" constructed in advance of the experiment. Each data plate is a data snapshot describing a tactical situation at a given time. The plate should include various information (course, speed, sensor reports, intelligence) available to a Tactical Action Officer to describe the situation fully.

d. Personnel

The personnel required for the experiment include observers of the output (threat assessors), a scenario controller, and a test director.

e. Support Program

In addition to scenario generating requirements, software to measure and record the times at which the messages are output and the times designated by the observers is required. This program, based on the ARPANET TYPESCRIPT feature, records all inputs and outputs to selected terminals.

3. General Context

a. Concept and Need

STAMMER2 is being developed to interact with a graphics display system to support the commander's decision processes. In order to assess the value of the technology we must determine its capabilities and shortcomings. For STAMMER2 to be useful to a decision maker it must be able to

evaluate threats by a comparison and collation of all-source reports and display the situations in a timely manner.

b. General Situation and Scenarios

A suitable number of data plates should be developed to vary the tactical situation significantly in both the complexity of the situations and the nature and number of the platforms involved. A data plate should be selected at random, its number recorded and the appropriately stored data input to the file which STAMMER2 will read. An observer, unaware of the contents of the data plate, monitors the contents of the display terminal, evaluates the information received, and describes the situations as he sees them.

4. Evaluation

a. Data collection

The following data should be collected for each trial: the data plate identification number, the situation assessment of the evaluator, and the situation assessment given by STAMMER2.

b. Analysis

The primary areas of concern for the stage-one experiments are: a comparison of the STAMMER2 output with the human assessments of the threats, and a determination of the times required for STAMMER2 to process different types of situations.

Post-test analysis will examine the data sheet to ascertain if STAMMER2 accurately evaluated or described and displayed all phase-one situations. This is compared with the evaluators' ability to do the same based on the same data.

c. Anticipated Results

It is anticipated that STAMMER2 will signal the existence of threats with very little time delay and that the situation and display will be presented accurately. For those cases where complex multiple situations are present simultaneously, the tests may indicate an overcrowding of the display and a lack of time to exercise the logic trace capability of STAMMER2.

5. Comments and Special Instructions

Each experimental subject will go through a short STAMMER2 indoctrination session during which he is given a summary of the theory and implementation of STAMMER2, a "threat" briefing to explain the scenario, and an overview of what is expected of him. Care should be taken to include in the data plates as many situations as possible that STAMMER2 has been designed to handle in realistic situations. It is recommended that several different operators be subjected to the complete set of plates and be asked to subjectively evaluate the displayed information to provide information which may be useful in the design of follow-on display experiments.

The stage-one experimentation discussed above will need to be carried out in the course of a full and careful evaluation of STAMMER2 as a tool in tactical assessment. For the purpose of this thesis, however, only the stage-two experiment (as discussed below) was actually performed.

B. STAGE-TWO EXPERIMENTS

1. Objective

The objectives of stage-two experiments are: to assess the technical capability of STAMMER2 actions on a dynamically changing data base generated by creating threat situations through the exercise of WIS (or its variants); to determine the "limits" of the capability of STAMMER2; and to obtain subjective opinions about STAMMER2.

2. Resources Required

a. Technical

The STAMMER2 program and associated computer hardware are required for this experiment. The Warfare Environment Simulator (WIS) program is available on the TENEX system at NOSC.

b. Display

An alphanumeric display system is required for players and the umpire team. The GENISCO system, a color graphic display, may be used to present the situation which is generated by WIS .

c. Scenario

An operationally realistic scenario is generated to test STAMMER2 in a tactical situation at a given time. The scenario should include various information (course, speed, sensor reports, intelligence) available to a Tactical Action Officer to describe the situation as it is normally presented by WES, via visual display and status boards. The evaluator will be allowed to interact with units under his "control" by ordering course and speed changes, and enabling various sensors available to him. The test director may insert various elements of intelligence as desired.

d. Personnel

The personnel required for the experiment include observers of the output (threat assessors), a scenario controller, and a test director.

e. Support Program

In addition to scenario generating requirements, software to measure and record the times at which the messages are output and the times designated by the observers is desirable. This program, based on the ARPANET TYPESCRIPT feature, records all inputs and outputs to selected terminals.

f. Evaluation questionnaire

Following a study of the measures of operational effectiveness which may be achieved in this interactive experiment, an evaluation questionnaire should be constructed to establish the level to which these measures have been satisfied.

3. General Context

a. Concept and Need

A complete technological assessment of the capability of STAMMER2 includes an evaluation in a realistic environment with a dynamically changing data base. The experiment must subject STAMMER2 to a wide of situations under varying levels of sensor activity in an effort to exercise all capabilities of STAMMER2 and try to probe the "limits" of those capabilities. These limits could be the size of data base, number of reports, or frequency of reports.

b. General Situation and Scenarios

WES scenarios will be generated to present a variety of sensor reports and levels of activity by platforms. An evaluator familiar with tactical war gaming will observe the play of the game while exercising STAMMER2 to the maximum extent possible.

Each experimental subject will go through a short STAMMER2 indoctrination session during which he will be given a summary of the theory and implementation of

STAMMER2, a 'threat' briefing to explain the scenario, and an overview of what is expected of him.

During the course of the actual experimental run, the evaluator may be assisted by the WIS operators and test director in the mechanism of input, output, and display to remove extraneous or distracting details of the game.

4. Evaluation

a. Data Collection

The following data should be collected for each scenario: the times at which the data base is updated, the situation analysis by STAMMER2, and the situation analysis by the operator.

These data will be augmented by the situation assessments made by the test director who, having become familiar with the full scenario, can assess the "ground truth" by having knowledge of the actions of all platforms throughout the scenario. The questionnaire given to each operator should consist of questions designed to assess the realism of the scenario and the utility of STAMMER2 as objectively as possible.

b. Analysis

Comparison is made between the situation descriptions by STAMMER2, the evaluators, and ground truth to assess the general accuracy of these descriptions. The questionnaire results may be presented in tabular form and the objective comments discussed in the conclusions.

c. Anticipated Results

While it may be possible to generate saturation points for STAMMER2, manifest in time delays wherein the STAMMER2 display significantly lags behind the WES scenario, the serial nature of report processing would seem to preclude this possibility, providing the query function of STAMMER 2 is not extensively used.

5. Comments and Special Instructions

The scenario tape should be maintained so that a given game may be replicated exactly with zero variance for different operators and so that the test director can thoroughly determine the situations as they exist (or are about to exist). Having the games on tape would allow trials without WES-trained operators to input the instructions required by a detailed script. In addition, an observer could monitor the Blue forces in one or more of the scenarios in a run without STAMMER2 and make comparisons with the exact run without STAMMER2. Finally, if the runs were not available on tape, very detailed scripts describing all of the actions of both the Blue and Orange forces might be written for each of the scenarios to accomplish the objectives of this test.

C. STAGE-THREE EXPERIMENTS

1. Objectives

The stage-three experiment has as its objectives to provide an operational evaluation of the military utility of STAMMER2 to a decision maker at an afloat task force command and control center in a simulated environment and to evaluate whether STAMMER2 improves the decision maker's ability to make rapid and accurate decisions. As in the case of the stage-one experiment, this level of experimentation are discussed here but were not performed as part of this research.

2. Resources Required

a. Technical

The STAMMER2 and WES programs and associated computer hardware are required for this experiment.

b. Display

An alphanumeric display system is required for each team of players and the umpire team. The GENISCO system, a color graphic display, may be used to present the situation which is generated by WES .

c. Scenario

An operationally realistic scenario is generated to test STAMMER2 in a controllable tactical situation. The scenario should include various information (course, speed, sensor reports, intelligence) available to a Tactical Action Officer to describe the

situation as it is normally presented by WES, via visual display and status boards. The teams will be allowed to interact with units under their command by ordering course and speed changes, enabling various sensors, and executing tactical orders against an opponent. The umpire may insert various elements of intelligence as desired.

d. Personnel

The personnel required for the experiment include observer teams to act as Blue and Orange Forces, a scenario controller, an umpire team, and a test director.

e. Support Program

In addition to scenario generating requirements, software to measure and record the times at which the messages are output and the times designated by the observers is required. This program, based on the ARPANET TYPESCRIPT feature, records all inputs and outputs to selected terminals.

f. Evaluation questionnaire

Following a study of the measures of operational effectiveness which may be achieved in this interactive experiment, an evaluation questionnaire should be constructed to establish the level to which these measures have been satisfied.

3. General Context

a. Concept and Need

It is possible that the value of STAMMER2 as a decision aid may only be assessed in an operational scenario. For example, it may enable a decision maker to better understand the threat situation which faces his forces and to make more accurate and timely decisions. These experiments would pit two opposing decision makers against each other in a simulated command and control environment and seek to measure the value of STAMMER2 by comparing the decision makers' performance operating with and without STAMMER2.

b. General Situation and Scenarios

The operational evaluation of STAMMER2 requires operationally realistic scenarios consisting of Blue and Orange task forces, which would be acted out with each team given the flexibility of exercising complete control over their forces as long as they do not countermand their ordered missions. For each scenario, an operational measure of effectiveness (MOE) or multiple MOEs should be selected and used to assess the operational utility of STAMMER2.

Because of the free-play flexibility given to the Blue and Orange force commanders, the war games and the resulting outcomes may vary significantly from trial to trial. Consequently, each scenario should be replicated in such a way as to avoid foreknowledge by a playing team. This

can be done by using a number of teams (based on the availability of assets), with multiple scenarios. By varying the use and non-use of STAMMER2 by a team (considering that a team is expected to play more than one game), a comparison of performance based on the MOEs may be made.

Each team should go through a brief indoctrination session during which the following points will be discussed: the theory and implementation of STAMMER2, a "threat" briefing to explain the scenario and an overview of the mission, and an overview of the measures of effectiveness.

4. Evaluation

a. Data Collection

The following data should be collected for each scenario: times at which the data base is updated, situation assessments by STAMMER2, situation assessments by the teams, actions taken by the decision maker, and data required to determine the selected operational measures of effectiveness.

These data may be augmented by the situation assessments made by the test director who, having become familiar with the full scenario, can assess the "ground truth" by having knowledge of the actions of all platforms throughout the scenario. The questionnaire given to each operator should consist of questions designed to assess

the realism of the scenario and the utility of STAMMER2 as objectively as possible.

b. Analysis

Data summaries may be made for each scenario and tabular displays should be made of the MOEs.

A correlation analysis should be made with the time sequence of actions taken by the decision maker based on the situation assessments made by STAMMER2 to try to determine how much the commanders utilize STAMMER2 and what information is most useful to them. A comparison of the game outcomes, as reflected by the MOEs, may be made using analysis of variance to see if there is any significant difference due to STAMMER2.

A comparison may be made of the game outcomes and the subjective appraisals of STAMMER2 to see if there was a relationship between how well the forces performed and how much they liked STAMMER2.

c. Anticipated Results

It is anticipated that there could be a large variance in the overall end- of-game operational measures of effectiveness. The possibility exists that the MOEs will be sufficiently ambiguous that any signal due to STAMMER2 may not be discernible. Thus, the subjective evaluations of the teams may provide the only discrimination in the test, with no quantitative assessment of the value of STAMMER2 possible.

The correlation analysis of the decision maker's actions and the STAMMER2 assessment should reveal the type of information that is most useful to a decision maker and how much he may grow to rely on STAMMER2. Overall, these experiments should yield important information about the acceptability of STAMMER2 to the decision maker and its utility to him.

5. Comments and Special Instructions

The usefulness of STAMMER2 to a decision maker may strongly depend on the manner in which STAMMER2 outputs are displayed to the Blue forces. This might be especially true in high density situations with a backlog of unprocessed sensor reports. The nature of the display dedicated to STAMMER2 should be considered prior to stage three experiments. The use of a terminal other than the graphic display terminal for query and response actions allows the picture to remain undisturbed until refinement of scale or sequential processing is requested. The display system is an integral part of STAMMER2 so the evaluation should be based on the most flexible display available.

A tape record of the game should be made of the inputs and outputs of each game so that evaluation personnel can reproduce a given trial and conduct a review of any point to assist in "what if" types of analysis.

The scenarios used should be operationally realistic, and the starting conditions and measures of effectiveness must be selected to exercise the full range of STAMMER2 capabilities and to yield useful information. The scenarios need not be developed just for the purpose of exercising STAMMER2. Instead, the Blue and Orange task forces should reflect as closely as possible, within the constraints of resources, the composition of forces currently available or projected as desired and the types of missions normally undertaken by each.

The subjects used in the experiments may be very important. For purposes of scientific credibility, the subjects must have command and control experience at high levels, preferably Navy Captains (O-6) or Admirals. A potential criticisms of experiments which attempt to determine the operational utility of a system is that the 'operators' did not reflect the potential user population. This would seem especially true when the evaluations depend heavily on subjective appraisals. There are practical difficulties in obtaining such test subjects. However, the issue is whether these tests can be viewed as scientific experiments or demonstrations without the operational realism engendered by their presence.

V. SAMPLE EXPERIMENT

A. DESCRIPTION

A sample experiment was selected from the preceding discussion to illustrate both the capability of STAMMER2 and and the practical application of these methods.

The stage-two experimental methodology was selected for this thesis. Selection criteria included the expertise of the test subjects available and the availability of the facilities at the Naval Postgraduate School. The test subjects selected were military officers with background in command and control concepts gained through formal education and a wide range of military experience. The Command, Control, and Communications Laboratory contains extensive graphic and text capability, and the STAMMER2 and WES software are available at the Naval Ocean Systems Center (NOSC), San Diego, California.

The Warfare Environment Simulator (WES) war game procedure involves the construction of a computer file which contains identification of all desired players, initial positions, courses, and speeds. Platforms may be selected from a resident data base which assigns sensor and weapons configurations to the units. Following game initialization, the units may be maneuvered on command from the player through a computer terminal, with display

of exact geographic location of known units available on selected graphic display monitors. Interaction of the units may be ordered by the players (as in attacks), but sensor reports are dependent on detection algorithms in WES. For example, detections by a surface search radar against another surface vessel occurs at 20 mile minimum separation. Sensor reports are given to the player via a status board text display and on the geographic display which is updated every game minute.

This experiment was not interactive in that the operator did not influence the series of events which were being generated by WES. It was, however, interactive in the sense that the operator had the capability to select different status board displays and change range scales as desired. The fully interactive war game as suggested in the third level of experiment would allow operator influence of the event stream. The prepared scenario used in this experiment, which included initial positions, initial courses and speeds, and orders to the various units are included in Appendix E.

The scenario for this experiment, generated in WES, consisted of three U.S. warships conducting a surface surveillance of a merchant lane in the vicinity of the North Atlantic, with neutral and (potentially) hostile platforms in the area. The warships, assisted by patrol aircraft conducting similar surveillance, were to locate,

track, and identify other units in the area. The experimental subject, "embarked" on the cruiser Belknap, had STAMMER2 available to assist in that mission. The STAMMER2 startup procedure is given in Appendix A.

The scenario was chosen based on the range of situations available and the relative realism of a non-combative confrontation. The scenario, tactics, sensors, and sensor characteristics are purely hypothetical. The data base [Appendix B] and rules [Appendix C] are assumptions tailored to this scenario, but are easily translatable to those which a Tactical Action Officer would recognize as having potential applications. The briefing of the subjects included an explanation of the artificialities.

A difficulty in the actual conduct of the experiment existed due to the nature of the reports on which STAMMER2 operates. The sensor reports received by an operator of WES are both visual cues on a graphics display (lines of bearing or new platform symbology), and in a table of detections (the Electronic Warfare Status Board, the Surface Status Board, or the Air Status Board). The form of the tabular report is formatted as to its numerical sequence, identification, and location (if applicable). The reports as generated by WES are not readable by STAMMER2. One example of the disparity is that STAMMER2 recognizes latitude and longitude in degrees and tenths rather than degrees and minutes. The LISP version of WES, LWES, does report in the

appropriate format, but the difficulty in using LWES, such as complex scenario generation and limited output capability, far outweigh this advantage.

The solution to the format disparity lies in the nature of WES. With duplicate platforms, starting at exact locations, given precisely the same order of instructions (such as course changes or sensor status), sensor reports are also duplicated. That is, the detections are based on maximum ranges of detection which are unchanged and which are combined in the WES detection algorithms in the same way, to produce duplicate reports at the same game time. This characteristic was explcited by running a game to a suitable point in time using the ARPANET typescript feature which copies all terminal interactions as they occur. This typescript shows the various reports received during a game, complete with the time of receipt. Synthesizing game reports, a complete file of reports [Appendix D] was built, in the format from which STAMMER2 could read its input data. Therefore, STAMMER2 was not receiving direct output from the sensors, but via an intermediate process not part of either WES or STAMMER2.

STAMMER2, using the TEKTRONIX 4014-1 terminal for both text and graphic display, was, in effect, a separate entity running concurrently with the game. The comparison of the subject's perspective of the situation and that of STAMMER2

is still valid only as long as simultaneous reporting is carefully maintained.

The artificiality of the message input could have biased the experiment to a remarkable degree but STAMMER2 has one feature which ameliorates this problem. STAMMER2 does not process the next sequential report in the file until the user quits the query mode of operation. That is, until the "quit" command is given there is no processing of the next sequential report. Because the experimental subject is not expected to be the actual operator of the STAMMER2 hardware, the operator can be instructed not to leave the query mode until the appropriate game time (which is displayed on the input/output terminal for WES).

Following a suitable period of game time where the subject is encouraged to exercise STAMMER2 a questionnaire was presented to gather the impressions and appraisal of the experiment in general and STAMMER2 in particular.

B. RESULTS AND CONCLUSIONS

The questionnaire supplied to test subjects, with a summary of results, is listed below. The intent of the questions was to verify the "realism" of the scenario and to gather a subjective evaluation of STAMMER2. The scoring scale of 1 to 12 was arbitrarily chosen, with 12 the most positive or most favorable response. The test

subjects consisted of eight Navy and Air Force officers, ranging in grade from O-3 to O-5, who are members of the Command, Control and Communications curriculum at the Naval Postgraduate School, Monterey, California. All have been exposed to decision theory and to an overview of the naval tactical environment. The Navy participants included three O-4s and one O-3. All but one of the Navy O-4s had Tactical Action Officer training.

STAMMER2 QUESTIONNAIRE

1. Prior to the scenario presentation, was the purpose of STAMMER2 adequately explained?

	1	2	3	4	5	6	7	8	9	10
Responses						1	1	3	2	1

Average 8.125

2. Was the scenario representative of your previous experience with WES or LWES?

	1	2	3	4	5	6	7	8	9	10
Responses					1	1		1	2	3

Average 8.375

3. Was the scenario realistic?

	1	2	3	4	5	6	7	8	9	10
Responses					1		1	1	3	2

Average 8.375

4. Were the displays consistent with the sensor reports?

	1	2	3	4	5	6	7	8	9	10
Responses							2	2	3	1

Average 8.375

5. Was sufficient time available to understand the sensor reports?

	1	2	3	4	5	6	7	8	9	10
Responses		1		1		1	2	1	2	1

Average 7.5

6. Were the STAMMER2 graphics clear?

	1	2	3	4	5	6	7	8	9	10
Responses	1		1	3		2		1		

Average 4.5

7. Was the STAMMER2 commentary understandable?

	1	2	3	4	5	6	7	8	9	10
Responses				1		3	1	2		1

Average 6.875

8. Were the STAMMER2 conclusions consistent with your conclusions?

	1	2	3	4	5	6	7	8	9	10
Responses						2	3	2	1	

Average 7.25

9. Did you use the explanation trace?

	1	2	3	4	5	6	7	8	9	10
Responses						1		3	3	1

Average 8.375

10. If the explanation trace was used, were the assertions consistent with the conclusions?

	1	2	3	4	5	6	7	8	9	10
Responses						2	3	1	2	

Average 7.375

11. Was the presence of STAMMER2 distracting?

	1	2	3	4	5	6	7	8	9	10
Responses	1	3				2			1	1

Average 4.75

12. Would an assistant who would filter information from STAMMER2 to the minimum essential level be useful?

	1	2	3	4	5	6	7	8	9	10
Responses					1	1			3	3

Average 8.5

13. Is STAMMER2 useful?

	1	2	3	4	5	6	7	8	9	10
Responses					1		1	4	2	

Average 7.75

The results show an overall trend towards an average response which could be considered favorable, in the vicinity of 8. The exceptions provide a clear contrast consistent with conditions noticed during the experiment.

Question 6 had the lowest average ranking. The difficulty encountered with the display may best be described by showing the display at three selected game minutes. Figure 1 is the last display without EW bearing lines, which occurs at minute 3. Figure 2 is the

display at game minute 14, which includes intelligence data (MIR1 and MIR2), surface contact reports, EW bearing lines and the location of friendly units. Figure 3 is the display at game minute 45 and is a compendium of all reports received to that point. Close observation of the screen as the graphics were being drawn showed a sequential entering of the information as it was being drawn. Some measure of order was discernible if this process was carefully followed, however the display is unreadable after it is completed.

A modification to the display package is required. One suggestion from the test subjects was that only the last two occurrences of a contact be displayed, with earlier reports available for historical review as directed by function key selection.

A benefit of STAMMER2 was demonstrated during the experiment. In every iteration of the display, intelligence information in the form of contact reports received from external sources was present on the screen as a reminder to the decision maker. Because a suspension concerning this information is created in memory, this information was available at all times for review and explanation as desired. This information was not subject to update in this scenario and did not seem to contribute to the general clutter discussed above.

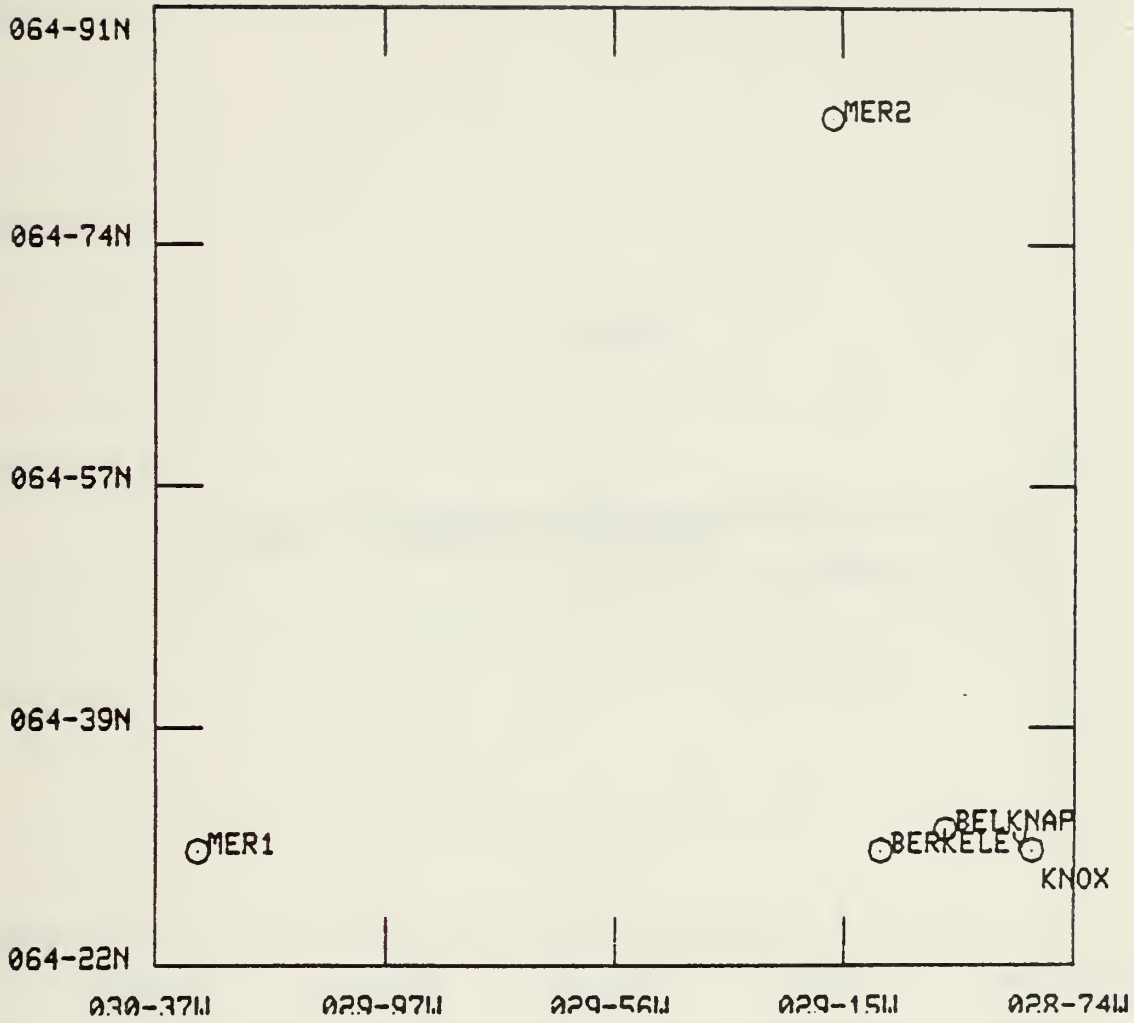

Fig. 1 STAMMER2 Graphics Display, Game Minute 3

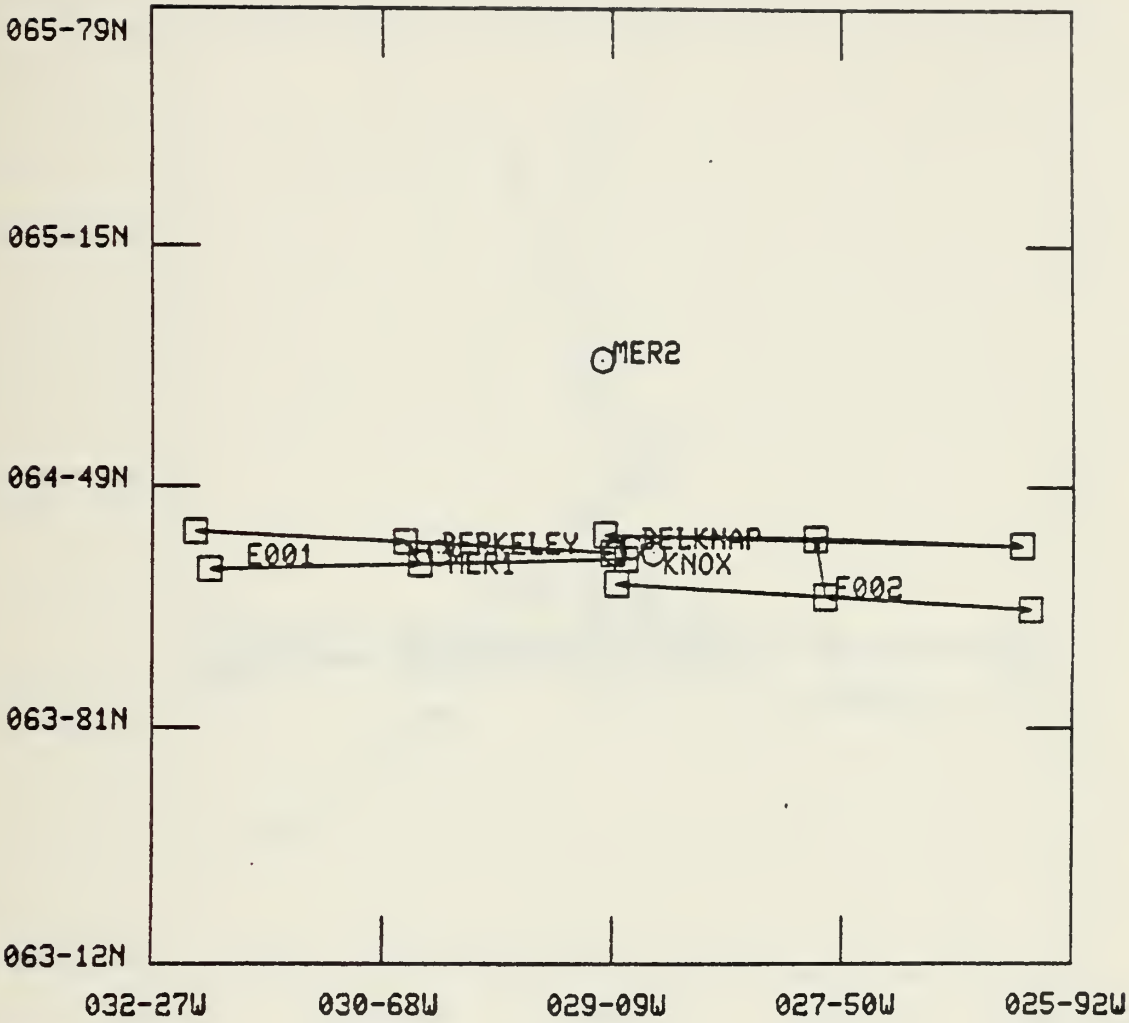

Fig. 2 STAMMER2 Graphics Display, Game Minute 14

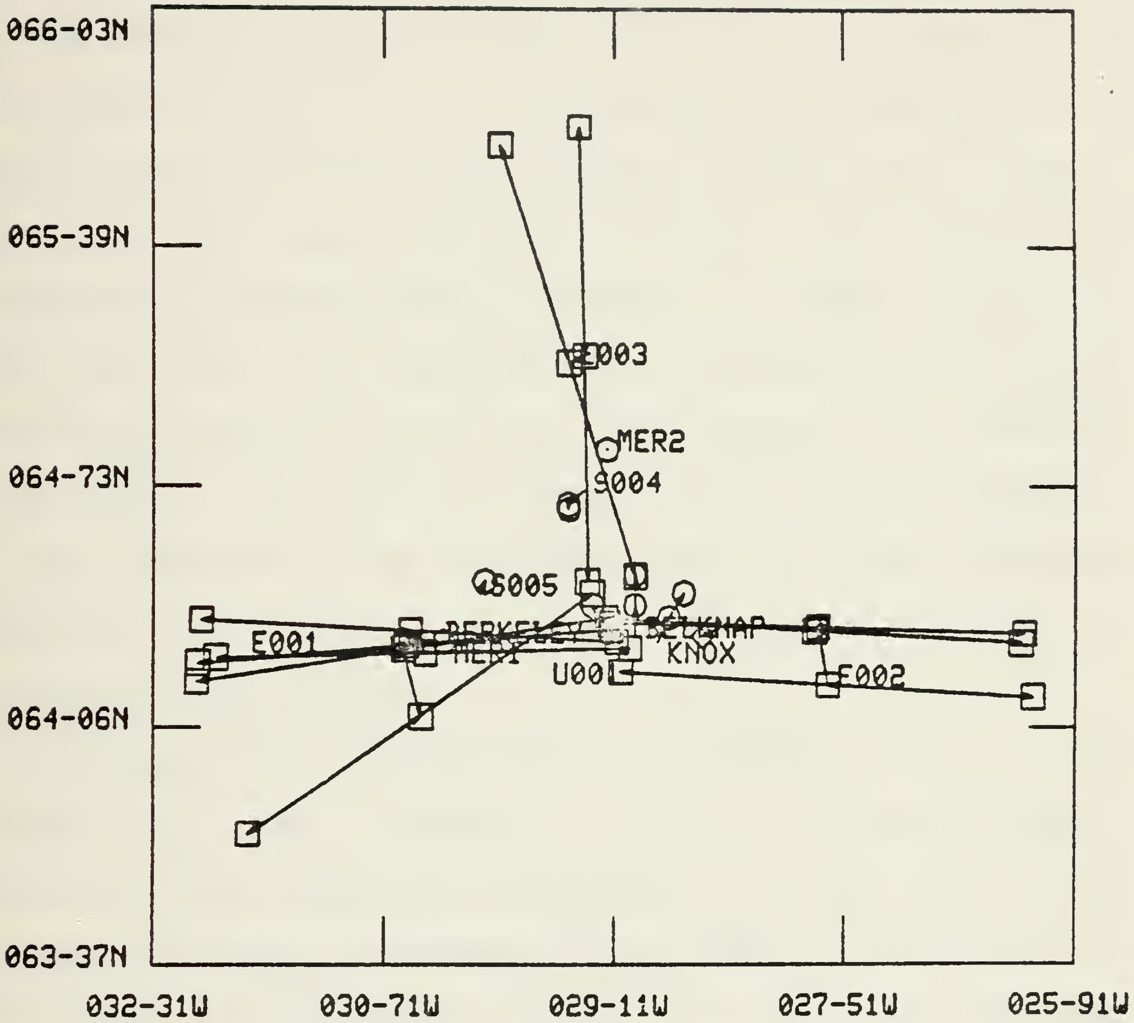

Fig. 3 STAMMER2 Graphics Display, Game Minute 45

There appeared to be a disparity between the STAMMER2 conclusions and those of the test subjects (questions 8 and 12). The difficulty was most apparent in cases where the test subjects were able to cross-fix passive detection bearing lines to a geographic position. Rather than conclude that an emitting platform exists at that point of intersection, STAMMER2 would continue to make conclusions based only on the individual EW reports. Here the major difficulty appeared to be in the construction of STAMMER2 rules rather than the functioning of the rules that exist. A rule correlating or position-fixing multiple passive detections of the same emitter would appear to be a relatively simple matter.

The explanation trace feature of STAMMER2 proved to be valuable as both a decision aid and as a training tool. While the users commented (question 9) that the trace included extraneous information, the level of understanding of the scenario increased markedly due to the increased appreciation of what rules were being applied and how STAMMER2 applied them.

The low average score for question 11 does not fit the general pattern of the questionnaire due to the wording of the question . The negative response indicates that STAMMER2 was not distracting to the user. The potential exists for the test subjects to became engrossed in the execution of STAMMER2 and the explanation trace to the point

where the conduct of the game becomes a minor consideration. This apparently was not the case here.

The remainder of the averages appear to be generally consistent with favorable response to both the scenario in general and to STAMMER2 in particular.

This experiment depends heavily on subjective evaluation. While the scenario and operation of STAMMER2 appear to have been convincing to the test subjects, there are no clear measures of operational effectiveness which can be quantified or analyzed. Given that the computer capacity is available to the location of the decision maker, STAMMER2 would appear to be a welcome addition to the repertoire of decision aids available to a commander. While this addresses the original objective namely the demonstration and evaluation of STAMMER2, it cannot fully answer the question of overall utility of the system to a decision maker in a real military environment.

Due to the experimental nature of STAMMER2, operator action was required for the message input procedure. The requirement to leave the explanation trace before the next message can be read is purely artificial and an accepted penalty in order to run the program simultaneously with WES. In an operational system, the mechanism for the transfer of data between a sensor and the decision maker should be understood and automatic. The addition of procedural steps such as function key or terminal inputs in

order to receive a sensor report may tend to be ignored during times of stress. In addition, this automatic transfer must be regulated to a suitable time period between reports. Reports from multiple sensors at short intervals would rapidly overload an operator required to view every report.

STAMMER2, as an investigation into the utility of artificial intelligence as applied to command and control, has progressed beyond the concept and technology demonstration phase of its existence. The somewhat artificial nature of the message transfer procedures is capable of approaching a "real world" condition in that machine and human readable formatted messages are in common use as in the RAINFORM reporting system. The production rules are similarly close to conditions which may in fact exist in an operational environment, although an agency, such as the Office of the Chief of Naval Operations, operating through the Surface Warfare Manpower and Training Requirements Division (OP-93), who could task the Fleet Training Centers (Atlantic and Pacific), could further refine them to reflect more realistic decision conditions.

A decision aid, even one directed toward a unique user (e.g., the Tactical Action Officer), may need to be distributed throughout a command center. The mass of data presented may require the decentralization of control to the

point where the TAC or staff watch officer function may be relegated to relatively discrete areas of responsibility. While the requirement for a prime decision maker exists, this role may well be one of command by negation. There is a time penalty for reevaluating previously screened information in a high density environment. Each decision maker would then require his own version of STAMMER2 that contains a tailored rule set. The present implementation of STAMMER2 already contains rules which may be useful and transportable to this practical application.

The benefit of this program aboard a ship, or at a major Fleet command center, for example, could be that a "super agency", consisting of intelligence and warfare specialists could provide a realistic assessment of situations which might be encountered and tailor propositions to reflect standardized intelligence, doctrine, and tactics in a pre-loaded database and system that would, in real time, be an "oracle" which would provide assistance and guidance to the decision maker. In case of questionable data the review capability of the process by which a decision was reached would assist the decision maker in logically analyzing his own thought process. The training benefit of the proposition review could increase the competence of the TAC who could see how the "panel of experts" approached the problem.

In summary, this thesis has presented background information concerning artificial intelligence and STAMMER2. Various testing methods were presented to evaluate the usefulness of STAMMER2 and an experiment was conducted and examined which used one of these methods. STAMMER2 appears to be a useful decision aid concept with great potential for further research and testing.

APPENDIX A

STAMMER2 GUIDE

The following is a procedural guide to beginning the STAMMER2 program for this experiment. After logging into the Naval Postgraduate School Command, Control, and Communications Laboratory UNIX system (DEC PDP 11/70), a TELNET link is established to the TOPS20 system at Naval Ocean Systems Center, San Diego, California, under account name NPS2. Following the prompt, the entries with asterisks are entered. Carriage returns (<cr>) are only required as indicated.

```
@ stammer2 <cr>
   _LOAD(TEMPHAK.COM)                                *
   compiled on 24-Feb-81 10:29:29
   FILE CREATED 24-Feb-81 10:28:55
   (DISPLOB redefined)
   (IN-LANE redefined)
   (RESULTPRINTER redefined)
   TEMPHAKCOMS
   (SENSORANGE reset)
   <NPS2>TEMPHAK.COM.1
   _(SETQ HOMESHIP (SETQ OWNSHIP 'BELKNAP))          *
   (OWNSHIP reset)
   (HOMESHIP reset)
   BELKNAP
   _(STAMMER)                                        *
   Welcome to version 2.5 of the STAMMER TSA system.
   Memory file? (Default is MEMORY.): MEM.JF <cr>    *
   Memory initialized.
   Rulefile? (Default is RULES.):RULES.JF <cr>       *
   Rules loaded
   What file would you like to take messages from?
   (Default is SCENE.ICE): SCENE.JF  <cr>            *
   Are you running on a Tektronix?no                 *
   Do you have a Tektronix available for display? no *
```

```
Passive detection. Heard SPS40 at bearing 258.29  Time: 12
Associated with track BERKELEY

   REPORT: BERKELEY was sighted in the merchant lane LANE1
   Question?  Quit
   Leaving EXPLAIN

Passive detection. Heard SPS39 at bearing 258.29  Time: 12
Associated with track BERKELEY

   REPORT: BERKELEY was sighted in the merchant lane LANE1
   Question?  BREAK
   (Explain broken)
:dribble]
```

APPENDIX B

DATA BASE

This is the data base from which STAMMER2 will draw its first assertions. It has been specifically tailored to the experiment scenario and does not reflect real-world conditions. STAMMER2 will update its data base as messages are received.

```
(MERCHANTLANE LANE1)
(LOCATION LANE1 ((63.33 -32.1)(65.9 -27.25)))
(OWNSHIP BELKNAP)
(ID BELKNAP FRIEND)
(ID-AMPLIFY BELKNAP MIL-BATTLE)
(ID BERKELEY FRIEND)
(ID-AMPLIFY BERKELEY MIL-BATTLE)
(ID KNOX FRIEND)
(ID-AMPLIFY KNOX MIL-BATTLE)
(ID U001 HOSTILE)
(ID-AMPLIFY U001 MIL-BATTLE)
(ID S004 FRIEND)
(ID-AMPLIFY S004 MIL-AUXIL)
(ID S005 FRIEND)
(ID-AMPLIFY S005 MIL-AUXIL)
(ID MER1 FRIEND)
(ID-AMPLIFY MER1 MIL-AUXIL)
(ID MER1 FRIEND)
(ID-AMPLIFY MER1 MIL-AUXIL)
(ID MER1 FRIEND)
(ID-AMPLIFY MER1 MIL-AUXIL)
(ID MER2 FRIEND)
(ID-AMPLIFY MER2 MIL-AUXIL)
(ID MER3 FRIEND)
(ID-AMPLIFY MER3 MIL-AUXIL)
(ID S008 HOSTILE)
(ID-AMPLIFY S008 MIL-BATTLE)
(ID S006 HOSTILE)
(ID-AMPLIFY S006 MIL-BATTLE)
(ID S007 HOSTILE)
(ID-AMPLIFY S007 MIL-BATTLE)
(ID HOST4 HOSTILE)
```

```
(ID-AMPLIFY HOST4 MIL-BATTLE)
STOP
```

APPENDIX C

RULES

The following are the production rules upon which STAMMER2 operates. As data from the data base or messages satisfy the conditions in each of the rules (labelled *<ITEM>), the data stream is built. When sufficient data exists, the rules fire and the appropriate actions are carried out. The plain text PRINFORM statement is the man-readable explanation of the assertion which has been built in STAMMER2.

```
INHERIT
(CONDITIONS ((ALIAS *PLAT *UNKNOWN)
             (TYPE *PLAT *TYP)
             (ID *PLAT *ID1)
             (ID-AMPLIFY *PLAT *IDMP)
             (CLASS *PLAT *CLS)
             (MEDIUM *PLAT *MED))
            ACTIONS
            ((TYPE *UNKNOWN *TYP)
             (ID *UNKNOWN *ID1)
             (ID-AMPLIFY *UNKNOWN *IDMP)
             (CLASS *UNKNOWN *CLS)
             (MEDIUM *UNKNOWN *MED))
            CONF 1.0 PRINFORM

"If  an unknown is identified, it inherits the properties of
its identification.")

KNOWN-PLAT
(CONDITIONS ((SIGHTING *P *S)
             (*UNLESS* (OWNSHIP *P))
             (FULLY-KNOWN *P)
             (ID-AMPLIFY *P *I))
            ACTIONS
            ((IDENTIFIED *P))
            CONF 1.0 PRINFORM
```

```
"Mark as identified if known with certainty")

NOT-LAST-SIGHTING-VER2
(CONDITIONS ((SIGHTING *PLAT *S1)
              (PREDECESSOR *S1 *S2)
              (*NOT* (SAME-AS *S2 NIL))
              (*UNLESS* (NOT-LAST *S2)))
             ACTIONS
             ((NOT-LAST *S2))
             CONF 1.0 PRINFORM

"If  a  sighting has a predecessor, then that predecessor is
not the last sighting.")

NOT-FIRST-SIGHTING
(CONDITIONS ((SIGHTING *PLAT *S1)
              (*UNLESS* (NOT-FIRST *S1))
              (PREDECESSOR *S1 *S2)
              (*NOT* (SAME-AS *S2 NIL)))
             ACTIONS
             ((NOT-FIRST *S1))
             CONF 1.0 PRINFORM

"If  an  earlier  sighting  occurs, record that the previous
sighting is not the first sighting.")

LAST-VIEW
(CONDITIONS ((SIGHTING *PLAT *S1)
              (*UNLESS* (NOT-LAST *S1)))
             ACTIONS
             ((LAST-SIGHTING *PLAT *S1))
             CONF .99 PRINFORM

"If  the  sighting is not followed, it is the last sighting.
(.99)")

FIRST-VIEW
(CONDITIONS ((SIGHTING *PLAT *S1)
              (*UNLESS* (NOT-FIRST *S1)))
             ACTIONS
             ((FIRST-SIGHTING *PLAT *S1))
             CONF .99 PRINFORM

"If  the sighting is not preceded, it is the first sighting.
(.99)")

SIMPLY-REACHABLE
(CONDITIONS ((CONTACT *CONT)
              (FIRST-SIGHTING *CONT *S1)
              (SIGHTING *PLAT *S2)
              (ID-AMPLIFY *PLAT MIL-BATTLE)
```

```
            (*NOT* (SAME-AS *CONT *PLAT))
            (*UNLESS* (OWNSHIP *PLAT))
            (POSITION *S1 *P1)
            (POSITION *S2 *P2)
            (TOS *S1 *T1)
            (TOS *S2 *T2)
            (SWR *P1 *T1 *P2 *T2))
           ACTIONS
           ((SIMPLY-WITHIN-REACH *S1 *S2))
           CONF .98 PRINFORM

"If  a  contact's  sighting  could  travel to a MIL-BATTLE's
sighting, then they are simply reachable. (.98)")

REACHABLE
(CONDITIONS ((CONTACT *CONT)
            (SIGHTING *CONT *S1)
            (SIGHTING *PLAT *S2)
            (*NOT* (SAME-AS *PLAT *CONT))
            (*UNLESS* (OWNSHIP *PLAT))
            (SIMPLY-WITHIN-REACH *S1 *S2)
            (*UNLESS* (BLOCKED-FROM *S1 *S2)))
           ACTIONS
           ((WITHIN-REACH *S1 *S2))
           CONF .97 PRINFORM

"If  two  sightings  are within reach of each other, and are
not blocked by patrol overflights, then are  reachable  from
each other. (.97)")

COULD-BE-COMBATANT
(CONDITIONS ((CONTACT *CONT)
            (FIRST-SIGHTING *CONT *S1)
            (ID-AMPLIFY *PLAT MIL-BATTLE)
            (*UNLESS* (OWNSHIP *PLAT))
            (LAST-SIGHTING *PLAT *S2)
            (WITHIN-REACH *S1 *S2))
           ACTIONS
           ((KNOWN-COMBATANT *CONT))
           CONF .15 PRINFORM

"If  a  contact's  position  could  be  reached  by  a known
combatant, then the contact might be a combatant (.15).")

NOT-KNOWN-COMBATANT
(CONDITIONS ((CONTACT *CONT)
            (*UNLESS* (KNOWN-COMBATANT *CONT)))
           ACTIONS
           ((TYPE *CONT MERCHANT))
           CONF .45 PRINFORM
```

```
"If a contact could not be any known combatant (as
determined by rule COULD-BE-COMBATANT), then it may be a
merchant (.45).")

POSS-RPT
(CONDITIONS ((PATROL *PTL)
             (CONTACT *CONT)
             (SIGHTING *CONT *S1)
             (SIGHTING *PLAT *S2)
             (ID-AMPLIFY *PLAT MIL-BATTLE)
             (*UNLESS* (OWNSHIP *PLAT))
             (SOURCE *S2 *PTL)
             (*NOT* (SAME-AS *S1 *S2))
             (*UNLESS* (DISSIMILAR *CONT *PLAT)))
            ACTIONS
            ((POSSIBLE-REPORT *CONT *PTL))
            CONF .95 PRINFORM

"If a patrol sights a MIL-BATTLE platform, and a contact is
similar to the platform, then the patrol report concerns the
contact. (.95)")

BLOCKER
(CONDITIONS ((CONTACT *CONT)
             (SIGHTING *CONT *S1)
             (SIGHTING *PLAT *S2)
             (ID-AMPLIFY *PLAT MIL-BATTLE)
             (*NOT* (SAME-AS *CONT *PLAT))
             (*UNLESS* (OWNSHIP *PLAT))
             (PATROL *PTL)
             (*UNLESS* (POSSIBLE-REPORT *CONT *PTL))
             (SIGHTING *PTL *S3)
             (NOT-LAST *S3)
             (SUCCESSOR *S3 *S4)
             (POSITION *S1 *P1)
             (POSITION *S2 *P2)
             (POSITION *S3 *P3)
             (POSITION *S4 *P4)
             (TOS *S1 *T1)
             (TOS *S2 *T2)
             (TOS *S3 *T3)
             (TOS *S4 *T4)
             (*OR* (CROSSPATHS *P1 *P2 *P3 *P4)
                   (GRAZE *P1 *P2 *P3 *P4))
             (*NOT* (WENT-BEFORE *P1 *T1 *P2 *T2 *P3 *T3 *P4
                         *T4))
             (*NOT* (WENT-AFTER *P1 *T1 *P2 *T2 *P3  *T3 *P4
                         *T4)))
            ACTIONS
            ((BLOCKED-FROM *S1 *S2))
            CONF .9 PRINFORM
```

```
"If  a  path between two sighting has not been detected by a
patrol, and it would have if they were sightings of the same
vessel, then they are different vessels. (.9)")

CREATEDETECT
(CONDITIONS ((SIGHTING *PLAT *SGT)
             (*UNLESS* (IDENTIFIED *PLAT))
             (*UNLESS* (DETECTION *PLAT))
             (SOURCE *SGT EW))
            ACTIONS
            ((DETECTION *PLAT))
            CONF 1.0 PRINFORM

"If the source of a sighting is EW, then mark it detected.")

CREATECONTACT
(CONDITIONS ((SIGHTING *PLAT *SGT)
             (*UNLESS* (IDENTIFIED *PLAT))
             (*UNLESS* (CONTACT *PLAT))
             (SOURCE *SGT RADAR))
            ACTIONS
            ((CONTACT *PLAT))
            CONF 1.0 PRINFORM

"If  radar  is the source of a platform's sighting, then the
platform is a contact.")

CREATEPLAT
(CONDITIONS ((SIGHTING *PLAT *SGT)
             (*UNLESS* (OWNSHIP *PLAT))
             (*UNLESS* (PLATFORM *PLAT)))
            ACTIONS
            ((PLATFORM *PLAT))
            CONF 1.0 PRINFORM

"Every sighting is a platform, except the ownship.")

SMALL-CRAFT9
(C0NDITIONS ((CONTACT *WHO)
             (FIRST-SIGHTING *WHO *S1)
             (SOURCE *S1 RADAR)
             (RANGE *S1 *R1)
             (LESS-THAN *R1 8)
             (STRENGTH *S1 STRONG))
            ACTIONS
            ((TYPE *WHO SUB)
             (MODE *WHO SURFACE))
            CONF .5 PRINFORM

"If  the range of a strong radar sighting is less than 3 nm,
```

```
and it is the first sighting, then the contact is possibly a
surface sub. (.5)")

SMALL-CRAFT6
(CONDITIONS ((CONTACT *X)
             (SIGHTING *X *SIGHT)
             (NOT-FIRST *SIGHT)
             (RANGE *SIGHT *R)
             (LESS-THAN *R 16)
             (GREATER-THAN *R 9)
             (STRENGTH *SIGHT WEAK)
             (SPEED *SIGHT *SPD)
             (*NOT* (GREATER-THAN *SPD 20)))
            ACTIONS
            ((*OR* (TYPE *X FISHING)
                   (TYPE *X PATROL)
                   (TYPE *X SUB)))
            CONF .15 PRINFORM

"If the range of a weak sighting is between 9 and 16, and
the speed is less than 20, then the contact is possibly a
sub or a patrol or a fishing vessel. (.15)")

SMALL-CRAFT5
(CONDITIONS ((CONTACT *WHO)
             (SIGHTING *WHO *S1)
             (NOT-FIRST *S1)
             (SOURCE *S1 RADAR)
             (RANGE *S1 *RANGE)
             (LESS-THAN *RANGE 16)
             (GREATER-THAN *RANGE 9)
             (STRENGTH *S1 WEAK)
             (SPEED *S1 *SPEED)
             (GREATER-THAN *SPEED 20))
            ACTIONS
            ((*OR* (TYPE *WHO SUB)
                   (TYPE *WHO PATROL)))
            CONF .3 PRINFORM

"If the range of a weak radar sighting is between 9 and 16,
and the speed is greater than 20, then the contact is
possibly a sub or a patrol. (.3)")

SMALL-CRAFT4
(CONDITIONS ((CONTACT *UNKNOWN)
             (SIGHTING *UNKNOWN *SIGHTING1)
             (LAND-DIST *SIGHTING1 *DIST)
             (SOURCE *SIGHTING1 RADAR)
             (RANGE *SIGHTING1 *RANGE)
             (LESS-THAN *RANGE 9)
             (GREATER-THAN *RANGE 3)
```

```
             (STRENGTH *SIGHTING1 WEAK)
             (LESS-THAN *DIST 50))
            ACTIONS
            ((*OR* (TYPE *UNKNOWN SUB)
                   (TYPE *UNKNOWN SHORE-PATROL)
                   (TYPE *UNKNOWN PLEASURE)
                   (TYPE *UNKNOWN COMMERCIAL)
                   (TYPE *UNKNOWN LANDING)))
            CONF .1 PRINFORM
```

"If the range of a weak sighting is between 3 and 9, and the distance from land of the sighting is less than 50, then the vessel may be a sub, a patrol, a pleasure craft, a landing craft, or a commercial craft. (.1)")

```
SMALL-CRAFT3
(CONDITIONS ((CONTACT *UNKNOWN)
             (SIGHTING *UNKNOWN *SIGHTING)
             (LAND-DIST *SIGHTING *DIST)
             (SOURCE *SIGHTING RADAR)
             (RANGE *SIGHTING *RANGE)
             (LESS-THAN *RANGE 9)
             (GREATER-THAN *RANGE 3)
             (STRENGTH *SIGHTING WEAK)
             (GREATER-THAN *DIST 50))
            ACTIONS
            ((TYPE *UNKNOWN SUB))
            CONF .35 PRINFORM
```

"If the range of a weak radar sighting is between 3 and 9, and the sighting is further than 50 miles from land, then the contact is a sub. (.35)")

```
SMALL-CRAFT2
(CONDITIONS ((CONTACT *UNKNOWN)
             (SIGHTING *UNKNOWN *SIGHTING)
             (NOT-FIRST *SIGHTING)
             (SOURCE *SIGHTING RADAR)
             (RANGE *SIGHTING *RANGE)
             (LESS-THAN *RANGE 3)
             (STRENGTH *SIGHTING WEAK)
             (SPEED *SIGHTING *SPEED)
             (*NOT* (GREATER-THAN *SPEED 3)))
            ACTIONS
            ((*OR* (TYPE *UNKNOWN DEBRIS)
                   (TYPE *UNKNOWN SUB)
                   (TYPE *UNKNOWN BUOY)))
            CONF .12 PRINFORM
```

"If the weak radar sighting is not known to be moving faster than 3 knots, then the contact is either a buoy, a sub, or

```
debris. (.12)")

SMALL-CRAFT1
(CONDITIONS ((CONTACT *UNKNOWN)
             (SIGHTING *UNKNOWN *SIGHTING)
             (NOT-FIRST *SIGHTING)
             (SOURCE *SIGHTING RADAR)
             (RANGE *SIGHTING *RANGE)
             (LESS-THAN *RANGE 3)
             (STRENGTH *SIGHTING WEAK)
             (SPEED *SIGHTING *SPEED)
             (GREATER-THAN *SPEED 3))
            ACTIONS
            ((TYPE *UNKNOWN SUB)
             (MODE *UNKNOWN PERISCOPE% OR% SNORKEL))
            CONF .6 PRINFORM

"If  the  weak  radar  sighting  is moving at greater than 3
knots, then the contact is a  sub  in  either  periscope  or
snorkel mode. (.6)")

ID-LANE
(CONDITIONS ((SIGHTING *SHIP *SIGETING)
             (MERCHANTLANE *LANE)
             (PLATFORM *SHIP)
             (LOCATION *LANE *LANELOC)
             (POSITION *SIGETING *POS)
             (IN-LANE *LANELOC *POS))
            ACTIONS
            ((INSIDE-A-MERCHANTLANE *SIGHTING)
             (*REPORT* *SHIP " was sighted in the merchant lane " *LANE)
            CONF 1.0 PRINFORM

"If  a ship is sighted within some merchantlane, then record
that it is inside that lane.")

INSIDE-A-STORM
(CONDITIONS ((STORM *STORM)
             (PLATFORM *SHIP)
             (*UNLESS* (IDENTIFIED *SHIP))
             (LOCATION *STORM *STMLOC)
             (SIGHTING *SHIP *SIGHTING)
             (POSITION *SIGETING *POS)
             (INSIDE *POS *STMLOC))
            ACTIONS
            ((TYPE *SHIP MERCHANT)
             (*REPORT* *SHIP " was sighted inside " *STORM))
            CONF -.25 PRINFORM

"If  a ship is sighted inside a storm, then the ship may not
be a merchant. (.25)")
```

```
CLOSE-POPUP
(CONDITIONS ((CONTACT *SHIP)
             (FIRST-SIGHTING *SHIP *SIGHTING)
             (RANGE *SIGHTING *RANGE)
             (LESS-THAN *RANGE 12))
            ACTIONS
            ((TYPE *SHIP MERCHANT))
            CONF -.2 PRINFORM

"If  the  first  sighting of a ship is within 12 nm, then it
may not be a merchant. (.2)")

DISTANT-POPUP
(CONDITIONS ((CONTACT *SHIP)
             (FIRST-SIGHTING *SHIP *SIGHTING)
             (RANGE *SIGHTING *RANGE)
             (GREATER-THAN *RANGE 30))
            ACTIONS
            ((TYPE *SHIP MERCHANT))
            CONF -.2 PRINFORM

"If  the  range of the first sighting is greater than 30 nm,
then the ship might not be a merchant.  (.2)")

COURSE-CHANGED
(CONDITIONS ((CONTACT *SHIP)
             (SIGHTING *SHIP *SIGHTING1)
             (NOT-FIRST *SIGHTING1)
             (NOT-LAST *SIGHTING1)
             (SUCCESSOR *SIGHTING1 *SIGHTING2)
             (COURSE *SIGHTING1 *COURSE1)
             (COURSE *SIGHTING2 *COURSE2)
             (*UNLESS*  (ROUGHLY-THE-SAME-COURSE-AS *COURSE1
                               *COURSE2)))
            ACTIONS
            ((TYPE *SHIP MERCHANT))
            CONF -.3 PRINFORM

"If  the course has changed significantly, then the sighting
may not be a merchant. (.3)")

SPEED-CHANGED
(CONDITIONS ((CONTACT *SHIP)
             (SIGHTING *SHIP *SIGHTING)
             (NOT-FIRST *SIGHTING)
             (NOT-LAST *SIGHTING)
             (SUCCESSOR *SIGHTING *SIGHTING2)
             (SPEED *SIGHTING *SPEED1)
             (SPEED *SIGHTING2 *SPEED2)
             (*UNLESS*   (ROUGHLY-THE-SAME-SPEED-AS  *SPEED1
```

```
                                *SPEED2)))
            ACTIONS
            ((TYPE *SHIP MERCHANT))
            CONF -.3 PRINFORM

"If  the  speed has changed significantly, then the sighting
may not be a merchant. (.3)")

FASTER-THAN-A-MERCHANT
(CONDITIONS ((CONTACT *SHIP)
             (SIGHTING *SHIP *SIGHTING)
             (NOT-FIRST *SIGHTING)
             (SPEED *SIGHTING *SPEED)
             (GREATER-THAN *SPEED 25))
            ACTIONS
            ((TYPE *SHIP MERCHANT))
            CONF -.25 PRINFORM

"If  the  speed  is  greater than 25 knots, then it is not a
merchant. (.25)")

SLOWER-THAN-A-MERCHANT
(CONDITIONS ((CONTACT *SHIP)
             (SIGHTING *SHIP *SIGHTING)
             (NOT-FIRST *SIGHTING)
             (SPEED *SIGHTING *SPEED)
             (LESS-THAN *SPEED 9))
            ACTIONS
            ((TYPE *SHIP MERCHANT))
            CONF -.15 PRINFORM

"If  the  speed  is  less than 9 knots, then it may not be a
merchant. (.15)")

MATCH-PLAT
(CONDITIONS ((FIRST-SIGHTING *PLAT1 *SGT1)
             (NOT-LAST *SGT1)
             (SUCCESSOR *SGT1 *SGT1S)
             (LAST-SIGHTING *PLAT2 *SGT2)
             (*UNLESS* (OWNSHIP *PLAT1))
             (*UNLESS* (OWNSHIP *PLAT2))
             (*NOT* (SAME-AS *PLAT1 *PLAT2))
             (POSITION *SGT1S *POS1S)
             (SPEED *SGT1S *SPD1)
             (POSITION *SGT1 *POS1)
             (COURSEFROM *POS1 *POS1S *CRS1)
             (TOS *SGT1 *T1)
             (POSITION *SGT2 *POS2)
             (TOS *SGT2 *T2)
             (LESS-THAN *T2 *T1)
             (COURSEFROM *POS2 *POS1 *CRS2)
```

```
                 (SPEEDFROM *POS2 *T2 *POS1 *T1 *SPD2)
                 (ROUGHLY-THE-SAME-COURSE-AS *CRS1 *CRS2)
                 (ROUGHLY-THE-SAME-SPEED-AS *SPD1 *SPD2))
                ACTIONS
                ((ALIAS *PLAT2 *PLAT1))
                CONF .5 PRINFORM

"If  the  course  and speed of two sightings are roughly the
same, and if one sighting's  position  would  be  the  other
sighting's extrapolated position, then the two sightings are
of the same vessel. (.5)")

OUTSIDE-ALL-LANES
(CONDITIONS ((SIGHTING *SHIP *SIGHTING)
                 (*UNLESS* (IDENTIFIED *SHIP))
                 (PLATFORM *SHIP)
                 (*UNLESS* (MEDIUM *SHIP AIR))
                 (*UNLESS* (INSIDE-A-MERCHANTLANE *SIGHTING)))
                 ACTIONS
                ((TYPE *SHIP MERCHANT))
                CONF -.08 PRINFORM

"If  a  sighting  is  outside  all  merchant lanes, then the
vessel might not be a merchant. (.08)")

STOP
```

APPENDIX D

MESSAGES

This is the message file which STAMMER2 will read, in sequential order (vice time order), during the execution of the scenario. It was built by putting WES generated reports into STAMMER2 readable form.

```
(BELKNAP 64.316 -28.97 0)
(BERKELEY RADAR 64.3 -29.085 03)
(KNOX RADAR 64.3 -28.616 03)
(MER1 RADAR 64.3 -30.3 0)
(MER2 RADAR 64.832 -29.167 0)
(E001 EW 270.0 K57 04)
(E001 EXTERNAL 274.0 BERKELEY 64.3 -29.10 05)
(E002 EW 92.0 K3 12)
(E002 EXTERNAL 90.0 BERKELEY 64.350 -29.15 14)
(E001 EXTERNAL 267.0 BERKELEY 64.333 -29.133 13)
(BELKNAP 64.4 -28.97 20)
(KNOX RADAR 64.367 -28.734 20)
(E001 EXTERNAL 263.0 BERKELEY 64.367 -29.167 20)
(E002 EXTERNAL 92.0 BERKELEY 64.367 -29.167 20)
(S004 EXTERNAL 64.716 -29.333 26)
(S004 EXTERNAL 64.680 -29.435 35)
(U001 EXTERNAL 64.350 -28.333 30)
(BELKNAP 64.485 -28.97 40)
(BERKELEY RADAR 64.4 -29.267 40)
(KNOX RADAR 64.433 -28.632 40)
(S005 EXTERNAL 64.433 -30.05 40)
(S004 EXTERNAL 64.567 -29.433 40)
(E003 EW 342.0 DONKA 41)
(E001 EXTERNAL 238.0 BERKELEY 64.433 -29.267 40)
(U001 EXTERNAL 64.316 -28.50 40)
(E003 EXTERNAL 359.0 BERKELEY 64.467 -29.30 48)
(S005 EXTERNAL 64.467 -30.016 45)
(S004 EXTERNAL 64.632 -29.467 45)
(S005 EXTERNAL 64.485 -29.984 50)
(S004 EXTERNAL 64.616 -29.50 50)
(U001 EXTERNAL 64.350 -28.648 50)
(U001 EXTERNAL 64.350 -28.648 50)
(E004 EXTERNAL 47.0 KNOX 64.5 -28.550 57)
(S006 EXTERNAL 64.848 -29.333 59)
(E004 EXTERNAL 47.0 KNOX 64.5 -28.534 59)
```

```
(BELKNAP 64.570 -28.97 60)
(BERKELEY RADAR 64.5 -29.367 60)
(KNOX RADAR 64.5 -28.534 60)
(E003 EXTERNAL 005.0 BERKELEY 64.534 -29.383 65)
(S005 EXTERNAL 64.534 -29.837 65)
(S004 EXTERNAL 64.570 -29.60 65)
(S006 EXTERNAL 64.832 -29.333 65)
(E004 EXTERNAL 48.0 KNOX 64.551 -28.485 70)
(E003 EXTERNAL 007.0 BERKELEY 64.570 -29.433 75)
(U001 EXTERNAL 64.367 -28.750 75)
(S007 EXTERNAL 64.766 -27.866 74)
(S005 EXTERNAL 64.570 -29.816 75)
(S008 EXTERNAL 64.267 -29.849 75)
(BELKNAP 64.532 -28.97 80)
(BERKELEY RADAR 64.57 -29.333 80)
(KNOX RADAR 64.586 -28.433 80)
```

APPENDIX E

SCENARIO

This is the scenario which will be generated by the Warfare Environmental Simulator (WES) program, resident in the TENEX system at NOSC, San Diego, California. This file sets initial ship identity, position, course and speed, and a basic file of orders to be carried out by each side. The orders, which represent contingency plans, are executed immediately following the initialization of WARGAM and the entry of PLAYER which is the user/interactive program for WES.

```
NORTH
YES
YES
SHIP
1.1 BERKL BERK
N00001 AAAA 900
64-18N  29-05W  330  15
SHIP
1.2 KNOX KNOX
N00002 BBBB 901
64-18N  28-50W  030  15
SHIP
1.3 BELKN  BELKN
N00003 CCCC 902
64-19N  28-55W  200  15
SHIP
7.1 VOROS  KRES2
N11111 EEEE 904
65-08N  29-30W  245  17
SHIP
7.2 SKORY  KASH
N11112 FFFF 905
65-12N  29-20W  180  20
SHIP
```

```
7.3 BURL  KOTLI
N11113 GGGG 926
65-05N  27-20W  217  18
SHIP
7.4 VAZNY  DDGPI
N11114 HHHH 927
64-22N  30-25W  140  17
SHIP
7.5 SSGN9  CHARL
N11115 IIII 928
64-20N  27-52W  275  25
SHIP
7.6 MER1  KAZBK
N22221 KKKK 920
64-18N  30-18W  220  16
SHIP
7.7 MER2  KAZBK
N22222 LLLL 921
64-50N  29-10W  215  16
SHIP
7.8 MER3  KAZBK
N22223 MMMM 922
64-18N  32-20W  038  17
ORDERS
BLUE PLAN ALFA
FOR BELKN REPORT ALL ALL TIME 1 999
FOR BERKL REPORT ALL ALL TIME 1 999
FOR KNOX REPORT ALL ALL TIME 1 999
PLACE A MARKER 64-10N 30-45W TIME 1 999
PLACE A MARKER 64-38N 29-30W TIME 1 999
PLACE A MARKER 65-18N 28-00W TIME 1 999

ORANGE PLAN ALFA
FOR SKORY REPORT ALL SURFACE TIME 1 999
FOR VAZNY COURSE 045 TIME 45
FOR BURL SPEED 20 TIME 65
FOR VOROS REPORT ENEMY AIR TIME 3  94
FOR BURL REPORT ALL ALL TIME 22 900
FOR VAZNY REPORT ALL SURFACE TIME 1 999
FOR SSGN9 REPORT ALL SURFACE TIME 10 24
FOR MER1 REPORT ALL SURFACE TIME 1 999


BYE
```

LIST OF REFERENCES

1. Winston, P. H., *Artificial Intelligence*, p. 1, Addison-Wesley, 1977.

2. Boden, M. A., *Artificial Intelligence and Natural Man*, p. 4, Basic Books, 1977.

3. George, F. H., *Models of Thinking*, p. 22, George Allen and Unwin, 1972.

4. Boden, op.cit.,p. 6.

5. Winston, op.cit., pp 11, 153, 225.

6. *Webster's New World Dictionary*, Vol. 1, p. 829, World Publishing, 1957.

7. Boden, op.cit.,p. 48.

8. Robinson, J. G.,"A New Tack for Tacticians", *Surface Warfare* , Vol. 5 No. 11, p. 24, November, 1982.

9. Naval Ocean Systems Center Technical Document 252, *STAMMER:System for Tactical Assessment of Multisource Messages, Even Radar*, by R. J. Bechtel and P. H. Morris, p. 4, May 1979.

10. U.S. Army Research Institute for the Behavioral and Social Sciences Research Report 1237, *Human Processes in Intelligence Analysis: Phase I Overview*, p. 2-6, December 1979.

11. Winston, op.cit, p. 241.

12. Naval Ocean Systems Center Technical Document 252, *STAMMER:System for Tactical Assessment of Multisource Messages, Even Radar*, by R. J. Bechtel and P. H. Morris, p. 32, May 1979.

13. Ibid., p. 36.

14. Naval Ocean Systems Center Technical Document 298, *STAMMER2 Production System for Tactical Situation Assessment, Volume I*, by D. C. McCall, P. H. Morris, D. F. Kibler, R. J. Bechtel,p. 41, October 1979.

15. Naval Ocean Systems Center Technical Document, Commander-in-Chief, Pacific Fleet, *Warfare Environment Simulator (WES) Users Guide*, 28 December 1979.

16. Naval Postgraduate School Technical Report NPS55-77-21, *Experimental Designs and Analyses for Initial ACCAT Test Bed Experimental Demonstrations*, by D. R. Barr, G. K. Poock, and F. R. Richards, p. 31, April 1977.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Technical Information Center Cameron Station Alexandria, Virginia 22314	2
2.	Superintendent ATTN: Code 0142 Naval Postgraduate School Monterey, California 93940	2
3.	Professor John M. Wozencraft, Code 74 Chairman, C3 Academic Group Naval Postgraduate School Monterey, California 93940	3
4.	Lt Col Jeffrey W. Johnson, USAF, Code 39 C3 Curricular Officer Naval Postgraduate School Monterey, California 93940	1
5.	Professor Donald R. Barr, Code 55Bn Department of Operations Research Naval Postgraduate School Monterey, California 93940	1
6.	Professor G. T. Howard, Code 55Hk Department of Operations Research Naval Postgraduate School Monterey, California 93940	1
7.	Commander ATTN: Code 8242 D.C. McCall Naval Ocean Systems Command San Diego, California 92152	6
8.	Commander, Cruiser-Destroyer Group Eight ATTN: LCdr John P. Ferranti, Jr. Fleet Post Office New York, New York 09501	2
9.	Strategic Air Command /INXY ATTN: Capt Greg Jay Offutt AFB, Nebraska 68113	1

10. Defense Advanced Research Projects Agency 1
Information Processing Techniques Office
ATTN: LCdr A. J. Dietzler
1400 Wilson Boulevard
Arlington, Virginia 22209

11. Commander 1
ATTN: Code 26221 P.L. LaRue
Naval Electronics Systems Command
Washington, D.C. 20360

192378

Thesis
F2643
c.2

Ferranti
Evaluation of the artificial intelligence program STAMMER 2 in the tactical situation assessment problem.

17 NOV 82	28223
FEB 17 85	30132
27 MAR 86	30818
3 MAY 86	30818
16 OCT 86	33348

192378

Thesis
F2643
c.2

Ferranti
Evaluation of the artificial intelligence program STAMMER 2 in the tactical situation assessment problem.

thesF2643
Evaluation of the artificial intelligenc
3 2768 001 01495 4
DUDLEY KNOX LIBRARY